MATHEMATICS

ON ELFVING'S PROBLEM OF IMBEDDING A TIME-DISCRETE MARKOV CHAIN IN A TIME-CONTINUOUS ONE FOR FINITELY MANY STATES. II

BY

C. L. SCHEFFER

(Communicated by Prof. J. POPKEN at the meeting of June 30, 1962)

1. The problem, considered in the preceding note (RUNNENBURG [1]), of finding a domain H_n to which the eigenvalues of a matrix of transition probabilities of a continuous parameter Markov chain must necessarily belong, is solved there by using a result in the corresponding problem for discrete parameter Markov chains, obtained by DMITRIEV and DYNKIN [2] and KARPELEWITSCH [3]. For a detailed statement of the problem we refer to [1].

In this note we want to show that the domain H_n can be found by using only the characterization of a Markov matrix given by (1) below.

We use the fact that a Markov-chain on a finite set consisting of n states $A_1, A_2, \ldots, A_n$, induces an abelian semi-group G of real linear transformations on an R_n: with the Markov matrix (p_{jk}) (where $p_{jk} = P\{\underline{x}_{n+1} = A_k | \underline{x}_n = A_j\}$) corresponds the linear transformation π defined by

$$\pi e_j = (p_{j1}, p_{j2}, \ldots, p_{in}),$$

where e_j is the j'th unit vector.

A necessary and sufficient condition that the matrix of a given linear transformation φ is a Markov matrix is

$$\varphi W \subset W, \tag{1}$$

where W is the set of all probability distributions on $\{A_1, A_2, \ldots, A_n\}$

$$W = \{x | x = (\xi_1, \xi_2, \ldots, \xi_n);\ V_j\, \xi_j \geqq 0;\ \sum_{j=1}^{n} \xi_j = 1\}. \tag{2}$$

2. Now let G be a continuous one-parameter abelian semi-group of linear transformations satisfying (1) (ι is the identical transformation on R_n)

$$G = \{\varphi_t | t \in (0, \infty);\ V_t{}^{(0,\infty)} \varphi_t W \subset W;\ V_t{}^{(0,\infty)} V_s{}^{(0,\infty)} \varphi_t \varphi_s = \varphi_{t+s};\ \lim_{t \downarrow 0} \varphi_t = \iota\}.$$

It is well-known (cf. e.g. FRÉCHET [4]) that under these conditions the eigenvalues of φ_t are $1, \lambda_1{}^t, \lambda_2{}^t, \ldots, \lambda_m{}^t$ $(m \leqq n-1)$, where $1, \lambda_1, \lambda_2, \ldots, \lambda_m$ are eigenvalues of φ_1, for some choice of the arguments of $\lambda_1, \ldots, \lambda_m$. Moreover there is at least one vector $p \in W$, which is invariant under G.

Let $\lambda = \varrho e^{i\psi}$ $(0 \leqq \varrho \leqq 1)$ be an eigenvalue of φ_1. We suppose that ψ is chosen in such a way that $\varrho^t e^{it\psi}$ is, for all $t > 0$, an eigenvalue of φ_t. We shall show below that

$$\text{(A)} \qquad \varrho \leqq e^{-\psi_0 \operatorname{tg} \pi/n}$$

holds, where

$$\psi_0 = \min\{\psi_0', 2\pi - \psi_0'\} \text{ and } \psi_0' = \psi - [\psi/2\pi]\, 2\pi,$$

i.e. ψ_0 is the minimum of all non-negative arguments of λ and $\bar{\lambda}$. It is easy to verify that the set of all complex numbers satisfying (A) is equal to the set H_n defined by (7) in [1]. Therefore (A) is equivalent to (6) in [1]. Furthermore (A) is trivially satisfied for $\psi = 0$ (for then $\psi_0 = 0$ also) and for $\varrho = 0$. Therefore we suppose henceforward $\psi \neq 0$ and $\varrho \neq 0$. As changing the unit of time amounts to multiplying ψ by a positive number we can always choose this unit in such a way that $\psi \neq 0$ entails $\lambda \notin R$ (R is the set of real numbers). Finally we assume $n \geqq 3$.

Let

$$z_0 = x_0 + i y_0,$$

where x_0 and y_0 are real vectors, be an eigenvector of φ_1 belonging to the eigenvalue λ. As $\lambda \notin R$, it follows that x_0 and y_0 are independent. Then the two-dimensional subspace V spanned by x_0 and y_0

$$V = \{u \mid u = \alpha x_0 + \beta y_0;\ \alpha, \beta \text{ real}\}$$

is invariant under G, and the movement induced by G in V is easily seen to be described by the equations

$$\text{(3)} \qquad \begin{cases} \varphi_t x_0 = \varrho^t (x_0 \cos t\psi - y_0 \sin t\psi), \\ \varphi_t y_0 = \varrho^t (x_0 \sin t\psi + y_0 \cos t\psi). \end{cases}$$

Introducing

$$\begin{aligned} W_0 &= \{x \mid x \in W;\ \exists_j \xi_j = 0\}, \\ U\ &= W \cap (p + V) \qquad , \\ U_0 &= W_0 \cap (p + V) \qquad , \end{aligned}$$

the condition (1) together with the invariance op p and V yields

$$\text{(4)} \qquad V_t^{(0,\infty)}\, \varphi_t\, U \subset U.$$

As U is a closed convex polygon [1]) in the plane $p + V$ and U_0 its boundary

---

[1]) It might seem that this statement is false for $n = 3$. However in that case our hypothesis $\lambda \notin R$ entails that 1 is the only real eigenvalue of φ_t (and hence of the adjoint transformation). Therefore $\{x \mid \xi_1 + \xi_2 + \xi_3 = 0\}$ is the only possible real invariant two-dimensional subspace and thus $V = \{x \mid \xi_1 + \xi_2 + \xi_3 = 0\}$, implying $W \subset p + V$, whence $U = W$ and $U_0 = W_0$.

the condition (4) is satisfied if and only if

$$V_t^{(0,\infty)}\ \varphi_t\, U_0 \subset U \tag{5}$$

is satisfied.

To derive a geometric condition equivalent with (5), consider a vector $w \in U_0$. The curve $\{\varphi_t w | t>0\}$ must, on account of (5), be in U. Hence a vector v_1 tangent to [1] $\{\varphi_t w | t \geqq 0\}$ in w must point from w to the interior of U. To make this condition more precise consider a vector v_2 pointing from w along U_0 in such a way that v_1 lies inside the angle [2] formed by $p-w$ and v_2.

If now we define

$$\theta_w = \sphericalangle\,(p-w, v_1);\ \theta_{0,w} = \sphericalangle\,(p-w, v_2), \tag{6}$$

then it is easily seen that (5) is satisfied if and only if

$$V_{w\in U_0}\ \theta_w \leqq \theta_{0,w} \tag{7}$$

is true.

3. It does not seem possible to deduce useful results directly from (7), as θ_w is in general a complicated expression depending on w. The only case in which θ_w is independent of w obtains when $|x_0|=|y_0|$ and the inner product $(x_0, y_0)=0$, which will be shown below. The general case can be reduced to this case by means of the following considerations:

Let σ be a real non-singular linear transformation on R_n. Then the semi-group

$$G^* = \sigma G \sigma^{-1}$$

consists of those linear transformations $\varphi_t{}^*$ for which

$$\varphi_t{}^* = \sigma\,\varphi_t\,\sigma^{-1}.$$

These $\varphi_t{}^*$ have the same eigenvalues as φ_t and eigenvectors

$$z^* = \sigma z,\ \text{with}\ z^* = x^* + iy^*;\ x^* = \sigma x;\ y^* = \sigma y,$$

where $z=x+iy$ is an eigenvector of φ_t.

Application of the transformation σ to $W, p, x_0, y_0, V, U, U_0, v_1, v_2$ yields $W^*, p^*, x_0{}^*, y_0{}^*, V^*, U^*, U_0{}^*, v_1{}^*, v_2{}^*$, i.e. $W^*=\sigma W$, etc.

We can now repeat the argument of 2 and arrive at equivalent conclusions. In particular, if we define

$$\theta_w^* = \sphericalangle\,(p^*-w, v_1{}^*);\ \theta_{0,w}^* = \sphericalangle\,(p^*-w, v_2{}^*), \tag{6*}$$

---

[1] Where $\varphi_0 \overset{\text{df}}{=} \iota$.

[2] There is a slight difficulty here on account of the fact that p may be in U_0. This case does not occur if $\psi \neq 0$, $\varrho \neq 0$: On account of (3) any vector $w-p$ performs a complete rotation around p in the time interval $(0,2\pi/\psi]$, and does not vanish in this time interval. Hence (5) and $p \in U_0$ are inconsistent.

then

$$(7^*) \qquad \bigvee_{w \in U_0^*} \theta_w^* \leqq \theta_{0,w}^*$$

is equivalent with (7).

In view of the remark at the beginning of this section, we choose σ in such a way that

$$(8) \qquad |x_0^*| = |y_0^*| \text{ and } (x_0^*, y_0^*) = 0.$$

Using (8) it is easy to calculate θ_w^*. For an arbitrary $w \in p^* + V^*$ we have

$$(9) \qquad w = p^* + \alpha(x_0^* \sin \gamma + y_0^* \cos \gamma) \qquad (\alpha, \gamma \text{ real}).$$

Consider $\{\varphi_t w | t \geqq 0\}$. A vector tangent to this curve in w is given by

$$(10) \qquad v_1^* = \lim_{h \downarrow 0} \frac{\varphi_h^* w - w}{h}.$$

Substitution of (9) and

$$(3^*) \qquad \begin{cases} \varphi_t^* x_0^* = \varrho^t(x_0^* \cos t\psi - y_0^* \sin t\psi) \\ \varphi_t^* y_0^* = \varrho^t(x_0^* \sin t\psi + y_0^* \cos t\psi) \end{cases}$$

in (10) gives after some calculation

$$(11) \qquad v_1^* = -\alpha A[x_0^* \sin(\gamma - \chi) + y_0^* \cos(\gamma - \chi)],$$

where

$$(12) \qquad A^2 = \psi^2 + \log^2 \varrho;\ 0 < \chi < \pi \text{ and } ctn\, \chi = -\psi^{-1} \log \varrho.$$

Then, using (8), (11) and (9), it follows that

$$(13) \qquad \cos \theta_w^* = \frac{(p^* - w, v_1^*)}{|p^* - w| \cdot |v_1^*|} = \cos \chi,$$

and hence (as θ_w^* and χ are both in $[0, \pi]$)

$$(14) \qquad \theta_w^* = \chi.$$

We thus see that θ_w^* is indeed independent of w. Our condition (7*) yields

$$(15) \qquad \chi \leqq \inf_{w \in U_0^*} \theta_{0,w}^*.$$

4. As σ is non-singular, W^* is like W an $(n-1)$-simplex. Hence U_0^*, being the boundary of the intersection of W^* with a plane that has at least one point (namely p^*) in common with W^*, is a convex polygon with at most n vertices, say $a_1, a_2, \ldots, a_n$. It is then easy to see, that the infinum in the right-hand side of (15) is reached if w is in one of the vertices of U_0^*. Therefore, if we introduce [1])

$$\left.\begin{aligned} \alpha_j &= \sphericalangle\, p^* a_j a_{j+1} \\ \alpha_j' &= \sphericalangle\, p^* a_j a_{j-1} \end{aligned}\right\} \quad (j = 1, 2, \ldots, n;\ \text{indices taken mod } n),$$

[1]) For any three vectors a, b and c we use the following notation:

$$\sphericalangle\, abc \overset{\text{df}}{=} \sphericalangle\, (a - b,\ c - b).$$

then

$$\inf_{w \in U_0^*} \theta^*_{0,w} = \min \alpha_j \text{ or } \min \alpha_j', \tag{16}$$

depending on whether the sense of rotation in $p^* + V^*$ given by p^*, w, v_1^* in that order is the same as the rotational sense given by $a_1, a_2, \ldots, a_n$ or that given by $a_n, a_{n-1}, \ldots, a_1$ respectively.

In **5** we shall give a proof of the fact that

$$\min_j \alpha_j \leqq \frac{n-2}{2n} \pi; \ \min_j \alpha_j' \leqq \frac{n-2}{2n} \pi, \tag{17}$$

with equality in at least one of these two relations if and only if the polygon under consideration is regular and p^* its centre.

Combination of (15), (16) and (17) yields

$$\chi \leqq \frac{n-2}{2n} \pi,$$

from which it follows that

$$\varrho \leqq e^{-\psi \operatorname{tg} \pi/n}. \tag{18}$$

Now ψ is of the form $\psi_0' + 2k\pi$ with $0 \leqq \psi_0' < 2\pi$ and k an integer. As $\bar{\lambda}$ is also an eigenvalue of φ_1, we may conclude

$$\exists_k \ \varrho \leqq e^{-(\psi_0' + 2k\pi) \operatorname{tg} \pi/n} \text{ and } \varrho \leqq e^{-(-\psi_0' - 2k\pi) \operatorname{tg} \pi/n}. \tag{19}$$

If we put

$$\psi_0 = \min \{\psi_0', 2\pi - \psi_0'\},$$

than either $\psi_0' + 2k\pi \geqq \psi_0$ or $-\psi_0' - 2k\pi \geqq \psi_0$, hence (A) holds for all eigenvalues λ of φ_1 for which $\psi \neq 0$, $\varrho \neq 0$, and therefore for all eigenvalues of φ_1 (cf. the remarks at the beginning of section 2).

5. It remains to prove (17). To this purpose suppose, with the notation of section **4**

$$\forall_j \ \alpha_j \geqq \frac{n-2}{2n} \pi. \tag{21}$$

Then it is possible to find vectors [1] $p_1 \in \langle p^*, a_1 \rangle$, $p_2 \in \langle p^*, a_2 \rangle$, ..., $p_n \in \langle p^*, a_n \rangle$ in such a way that

$$\sphericalangle\, p^* a_n p_1 = \sphericalangle\, p^* p_1 p_2 = \ldots = \sphericalangle\, p^* p_{n-1} p_n = \frac{n-2}{2n} \pi.$$

Then, as $p_n \in \langle p^* a_n \rangle$

$$\frac{|p_n - p^*|}{|a_n - p^*|} \leqq 1,$$

[1]) We denote by $\langle b_1, b_2, \ldots, b_k \rangle$ the closed convex hull of $\{b_1, \ldots, b_k\}$:

$$\langle b_1, b_2, \ldots, b_k \rangle = \{x | x = \alpha_1 b_1 + \alpha_2 b_2 + \ldots + \alpha_k b_k; \ \forall_j \alpha_j \geqq 0; \ \sum_{j=1}^{k} \alpha_j = 1\}.$$

with equality if and only if $\forall_j p_j = a_j$. Introducing $\alpha_j'' = \sphericalangle\, p^* p_j p_{j-1}$ $(j=2, \ldots, n)$ and $\alpha_1'' = \sphericalangle\, p^* p_1 a_n$, we find

$$1 \geqq \frac{|p_n - p^*|}{|a_n - p^*|} = \frac{|p_n - p^*|}{|p_{n-1} - p^*|} \cdot \frac{|p_{n-1} - p^*|}{|p_{n-2} - p^*|} \cdots \frac{|p_1 - p^*|}{|a_n - p^*|} = \prod_{=1}^{n} \frac{\sin \dfrac{n-2}{2n}\pi}{\sin \alpha_j''}.$$

Hence

$$\prod_{j=1}^{n} \sin \alpha_j'' \geqq \sin^n \frac{n-2}{2n} \pi, \tag{22}$$

with equality if and only if $\forall_j \alpha_j'' = \alpha_j'$.

On the other hand, as the sum of the angles $(n\pi)$ of the n triangles $\langle a_n, p_1, p^* \rangle, \langle p_1, p_2, p^* \rangle, \ldots, \langle p_{n-1}, p_n, p^* \rangle$ is equal to

$$\sum_{=1}^{n} \alpha_j'' + 2\pi + n \cdot \frac{n-2}{2n} \pi,$$

we have

$$\sum_{j=1}^{n} \alpha_j'' = n\pi - 2\pi - n \cdot \frac{n-2}{2n} \pi = \frac{n-2}{2} \pi.$$

As $-\log \sin x$ is a convex function for $0 < x < \pi$, we have

$$\prod_{j=1}^{n} \sin \alpha_j'' \leqq \sin^n \frac{1}{n} \sum_{j=1}^{n} \alpha_j'' = \sin^n \frac{n-2}{2n} \pi \tag{23}$$

with equality if and only if

$$\forall_j\, \alpha_j'' = \frac{n-2}{2n} \pi.$$

Combination of (22) and (23) yields:

$$\forall_j\, \alpha_j' = \alpha_j'' = \frac{n-2}{2n} \pi. \tag{24}$$

We can now repeat this argument, starting from

$$\forall_j\, \alpha_j' \geqq \frac{n-2}{2n} \pi, \tag{25}$$

which follows from (24). We then arrive at the conclusion

$$\forall_j\, \alpha_j = \frac{n-2}{2} \pi. \tag{26}$$

Hence, either of (21) or (25) entails (24) and (26). In other words: Either

$$\forall_j\, \alpha_j = \alpha_j' = \frac{n-2}{2n} \pi,$$

in which case $\langle a_1, a_2, \ldots, a_n \rangle$ is regular with centre p^*, or

$$\min_j \alpha_j < \frac{n-2}{2n}\pi \text{ and } \min_j \alpha_j' < \frac{n-2}{2n}\pi,$$

which proves (17).

REFERENCES

1. RUNNENBURG, J. TH., On Elfving's problem of imbedding a time-discrete Markov-chain in a time-continuous one for finitely many states I, These Proceedings 536.
2. DMITRIEV, N. and E. DYNKIN, On characteristic roots of stochastic matrices, Isvestija, ser. mat., **10**, 167–194 (1946).
3. KARPELEWITSCH, F. I., On the characteristic roots of a matrix with nonnegative elements, Isvestija, ser. mat., **15**, 361–383 (1951).
4. FRECHET, M., Méthode des fonctions arbitraires, Gauthier–Villars, Paris (1952).